

J02275. Minute virus of m...[gi:332293]

LOCUS MVMPG 5149 bp ss-DNA linear VRL 22-MAY-1995

DEFINITION Minute virus of mice, complete genome.

ACCESSION J02275 M12520 M12521 M14704

VERSION J02275.1 GI:332293

KEYWORDS alternative splicing; capsid protein; complete genome; nonstructural protein.

SOURCE Mice minute virus

ORGANISM Mice minute virus

Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

REFERENCE 1 (bases 1 to 5149)

AUTHORS Astell,C.R., Thomson,M., Merchlinsky,M. and Ward,D.C.

TITLE The complete DNA sequence of minute virus of mice, an autonomous parvovirus

JOURNAL Nucleic Acids Res. 11 (4), 999-1018 (1983)

MEDLINE 83143341

PUBMED 6298737

REFERENCE 2 (bases 1 to 5149)

AUTHORS Astell,C.R., Gardiner,E.M. and Tattersall,P.

TITLE DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and comparison with the DNA sequence of the fibrotropic prototype strain

JOURNAL J. Virol. 57 (2), 656-669 (1986)

MEDLINE 86115415

PUBMED 3502703

REFERENCE 3 (sites)

AUTHORS Morgan,W.R. and Ward,D.C.

TITLE Three splicing patterns are used to excise the small intron common to all minute virus of mice RNAs

JOURNAL J. Virol. 60 (3), 1170-1174 (1986)

MEDLINE 87061199

PUBMED 3783817

COMMENT Original source text: Minute virus of mice (strain MVM(p)), passed in mouse 1 (variant A-9) cells.

The parvoviridae family contains two groups that infect mammalian hosts: (i) defective (helper-dependent) adeno-associated viruses, and (ii) autonomous (helper-independent) parvoviruses. MVM is a member of the latter group. Both groups have been demonstrated to package both plus and minus strands (in separate particles) of the ss-DNA genome, though the minus strand is more typically packaged in the latter group.

The sequence below corresponds to the plus (+) strand, also

referred to as the C-strand. The minus (-) strand is also referred to as the V-strand.

The 3' and 5' termini both exhibit the potential for forming stable 'fold-back' hairpins; these sequences appear to play a role in replication [1].

revision 4804 4870 a-65bp-a in [2]; aa in [1] [2]
revises [1].

ORIGIN 5' end of genome; 415 bp upstream of PstI site.

```

1 attttagaa ctgaccaacc atgtcacgt aagtgcgtg atgacgcgcg ctgcgcgc
61 gcctcgac gtcacacgtc acttacgtt cacatgggt gtcagttcta aaaatgataa
121 gcggttcagg gagttaaac caaggcgcga aaaggaagtg ggcgtggtt aaagtatata
181 agcaactact gaagtcagtt acttatctt tcttcattc tgtgagtcga gacgcacaga
241 aagagagtaa ccaactaacc atggctggaa atgcttactc tgatgaagtt ttggagcaa
301 ccaactggtt aaaggaaaaa agtaaccagg aagtgttctc atttgtttt aaaaatgaaa
361 atgtcaact gaatggaaaa gatatcggtt ggaatagttt caaaaaagag ctgcaggagg
421 acgagctgaa atcttacaa cgaggagcgg aaactacttg ggaccaaagc gaggacatgg
481 aatgggaaac cacagtggat gaaatgacca aaaagcaagt attcattttt gattcttgg
541 taaaaaatg ttatttcaa gtgcctaaca caaagaatat attccctggt gatgttaatt
601 ggttgtgca acatgaatgg ggaaaagacc aaggctggca ctgccatgt ctaattggag
661 gaaaggactt tagtcaagct caagggaaat ggtggagaag gcaactaaat gtttactgga
721 gcagatgggtt ggtaacagcc tgaatgtgc aactaacacc agctgaaaga attaaactaa
781 gagaatagc agaagacaat gagtgggtt ctctacttac ttataagcat aagcaaacc
841 aaaaagacta tccaagtgt gttttttt gaaacatgtat tgcttactat ttttaacta
901 aaaaagaaaat aagcactagt ccaccaagag acggaggcta tttcttagc agtgactctg
961 gctggaaaac taactttta aaagaaggcg agcgcctatct agtgagcaaa ctatacactg
1021 atgacatgca gccagaaacg gttgaaacca cagtaaccac tgcgcaggaa actaagcgc
1081 gcagaattca aactaaaaaa gaagttctt taaaactac actaaagag ctggcgcata
1141 aaagagtaac ctcaccagag gactggatga tgatgcagcc agacagttac attgaaatga
1201 tggctcaacc aggtggagaa aacctgctga aaaatacgct agagatttg acactaactc
1261 tagccagaac caaaacagca tttgacttaa ttttagaaaa agctgaaacc agcaaactaa
1321 ccaactttc actgcctgac acaagaacct gcagaatttt tgctttcat ggctggaaact
1381 atgttaagt ttgcctatgtt atttgctgt ttttaacag acaaggaggc aaaagaaata
1441 ctgtttatt tcatggacca gccagcacag gcaaactat tattgcacaa gccatagcac
1501 aagcagttgg caatgttgg tgctataatg cagccatgt aaactttcca ttaatgact
1561 gtaccaacaa gaacttgatt tggtagaaag aagctggtaa cttggacag caagtaaacc
1621 agtttaagc catttgctct ggtcaaacta ttgcattga tcaaaaagga aaggcagca
1681 aacagattga accaacacca gtcatcatga ccacaaatga gaacattaca gtggcagaa
1741 taggctgcga agaaagacca gaacacactc aaccaatcag agacagaatg cttacattc
1801 atctaacaca tacctgcct ggtgactttt gtttgggttga caaaaatgaa tggcccatga
1861 tttgtgcttg gttggtaaag aatggttacc aatctaccat ggcaagctac tttgtctaaat
1921 gggccaaagt tcctgattgg tcagaaaaact gggcggagcc aaaggtgcca actcctataa
1981 atttactagg ttccgcacgc tcaccatca cgacacccaa aagtacgcct ctcagccaga
2041 actatgact aactccactt gcatcgatc tcgaggaccc ggccttagag ccttggagca

```

2101 caccaaatac tcctgttgcg ggcactgcag aaacccagaa cactggggaa gctggttcca
2161 aaggcctgcca agatggtcaa ctgagccaa ctgggtcaga gatcgaggag gatttgagag
2221 cgtgcgttcgg tgcggAACCG ttgaagaaag acttcagcga gccgctgaac ttggactaag
2281 gtacgatggc gcctccagct aaaagagcta aaagaggtaa gggttaagg gatggtttgt
2341 tgggggtta ttatgttta attacctgtt ttacaggcct gaaatcacctt gggttaggt
2401 tgggtgcctc ctggctacaa gtacctggga ccagggaaaca gccttgacca aggagaacca
2461 accaatccat ctgacggcgc tgccaaagag cagcagcagg cctatgtca atacatcaaa
2521 tctggaaaaaa atccttacct gtacttctc gctgctgatc aacgccttat tgaccaaacc
2581 aaggacgcca aagactgggg aggcaagggtt ggtcactact tttttagaac caagcgcgc
2641 ttgcaccta agcttgcac tgactctgaa cctggaacctt ctgggttaag cagagctgg
2701 aaacgcacta gaccacctgc ttacattttt attaaccagg ccagagctaa aaaaaaactt
2761 acttctctg ctgcacagca aagcgtcaa accatgagtg atggcaccag ccaacctgac
2821 agcggaaacg ctgtccactc agctgcaaga gttgaacgg cagctgacgg ccctggaggc
2881 tctgggggtg ggggctctgg cgggggtggg gttgggtttt ctactgggtc ttatgataat
2941 caaacgcatt atagattctt gggtgacggc tgggttagaaa ttactgcact agcaactaga
3001 ctgtacatt taaacatgcc taaatcgaa aactattgca gaatcagagt tcacaataca
3061 acagacacat cagtcaaagg caacatggca aaagatgtg ctcacatggca aatttggaca
3121 ccatggagct tggtgatgc taatgcttgg ggagttggc tccagccaag tgactggcaa
3181 tacatttgc acaccatgag ccagcttaac ttggtatcac ttgatcaaga aatattcaat
3241 gtatgtcga aaactgttac agagcaagac ttaggaggc aagctataaa aatatacaac
3301 aatgaccta cagctgcat gatgggtgca ttagactcaa acaacattt gcccatacaca
3361 cctgcagcaa actcaatgga aacacttggt ttctacccct gggaaaccaac catagcatca
3421 ccatacaggta actattttg cgttgacaga gatcttgc tgacctacga aatcaagaa
3481 ggcacagttg aacataatgt gatgggaaca cccaaaggaa tgaattctca attttttacc
3541 attgagaaca cacaacaat cacattgctc agaacagggg acgaatttgc cacaggtaact
3601 tactactttg acacaaattc agttaaactc acacacacgt ggcaaaacca ccgtcaactt
3661 ggacagcctc cactgctgac aaccttcctt gaagctgaca ctgatgcagg tacacttact
3721 gctcaaggga gcagacatgg aacaacacaa atgggggtta actgggttag tgaagcaatc
3781 agaaccagac ctgctcaagt aggatttgc caaccacaca atgacttgc agccagcaga
3841 gctggaccat ttgctcccccaaaaatccca gcagatatta ctcaaggagt agacaaagaa
3901 gccaatggca gtgttagata cagttatggc aaacagcatg gtggaaatttggc ttacatccat
3961 ggaccagcac cagagcgcta cacatggat gaaacaagct ttgggtcagg tagagacacc
4021 aaagatggtt ttatcaatc agcaccacta gttgtccac caccactaa tggcatttt
4081 acaaatgcaaa accctattgg gactaaaaat gacattcatt ttcaatgt tttaacagc
4141 tatggtccac taactgcatt ttcaacccca agtccctgtat accctcaagg acaaataatgg
4201 gacaaagaac tagatcttgc acacaaacct agacttcaca taactgctcc atttgtttgt
4261 aaaaacaatg cacctggaca aatgttggtt agataggac caaacctaact tgaccaatatt
4321 gatccaaacg gagccacact ttctagaatt gttacatacg gtatctttt ctggaaagga
4381 aaactaacca tgagagcaaa acttagagct aacaccactt ggaacccagt gtaccaagta
4441 agtgctgaag acaatggcaaa ctcatacatg agtctaacta aatggttacc aactgctact
4501 gggaaacatgc agtctgtgcc gcttataaca agacctgttgc ctggaaatac ttactaacta
4561 accatgctttt ttctttctgtt acttcataata ttatggatc taataaagat acaacataga
4621 aatataatat tacgtataga ttatggaaat agaataatg ggtacttagt aactgtttaaa

4681 aataatagaa ccttggaat aacaagatag ttagttggtt aatgttagat agaataagaa
4741 gatcatgtat aatgaataaa agggtggaag ggtggttggt aggttaatgt tagatagaat
4801 aagaagatca tgtataatga ataaaagggt ggaagggtgg ttggtaggtt ttcccttaga
4861 ctgatgtta aggaccaaaa aaataataaa actttttaa aactcaacca agactactgt
4921 ctattcagtg aaccaactga accattagta ttactatgtt tttagggtgg gagggtggga
4981 gatacatgtg ttgcgttatga gcgaactggt actgggtgg tgctctgctc aaccaaccag
5041 accggcaaag ccggcttgtt tggtagcg caaccaacca gtaccagttc gctcatagcg
5101 aacacatgtt tctcccaccc tcccacccta aaaacatagt aatactaataat

NC_004713. LuIII virus, comp...[gi:29742044]

LOCUS NC_004713 5135 bp ss-DNA linear VRL 20-AUG-2003

DEFINITION LuIII virus, complete genome.

ACCESSION NC_004713

VERSION NC_004713.1 GI:29742044

KEYWORDS

SOURCE LuIII virus (LuIV)

ORGANISM LuIII virus

Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

REFERENCE 1 (bases 1 to 5135)

AUTHORS Diffoot,N., Chen,K.C., Bates,R.C. and Lederman,M.

TITLE The complete nucleotide sequence of parvovirus LuIII and
localization of a unique sequence possibly responsible for its
encapsidation pattern

JOURNAL Virology 192 (1), 339-345 (1993)

MEDLINE 93297126

PUBMED 8517025

COMMENT REVIEWED REFSEQ: This record has been curated by NCBI staff. The
reference sequence was derived from M81888.

Coding regions were annotated at the NCBI following the annotation
of closely related Mouse parvovirus 1 (U12469).

1 atcattttta gaactaacca accatgttca cgtaagtgcgt gttggatgcgc ggcgtacgcg
61 cgctgccttc ggcagtcaca cgtcacttac gtctcacatg gttggtagt tctaaaaatg
121 ataagcggtt cagggagtt aaaccaaggc gcgaaaagga agtgggcgtg gtttaagta
181 tataagcgcac acgttaagtc agttacttac tcttcgcctt attctgttaag tcgagacaca
241 cagagtaacc aactaaccaa ctggccatgg ctggaaacgc gtactctgat gaagtttg
301 gaacaactaa ctgggtgaag gataagagca accaggaagt attctcattt gtttttaaaa
361 atgaggatgt tcagctcaat ggaaaaata tcggatggaa cagttacaga aaggagctgc
421 aagaggagga gctgaaatct ttacaacgag gagctgaaac tacctggac cagagcgagg
481 acatggaatg ggaatcttca gtggatgaac tgaccaaaaa gcaagtattt attttgact
541 cttagttaa aaagtgtctc ttgttgttac tgaggcacaaa gaacatagct ccttagtgcgt
601 ttacttgggtt tgtacagcat gaatggggaa aagaccaagg ctggcactgt catgtgccta
661 ttggaggcaaa gaactttagc caggctcaag gaaaatgggtt gaggagacaa ttaaatgttt
721 actggatgtt atgggttgtt acagcctgtt gcgtgcgtt atcaccatgtt gaaagaattt
781 aactaagaga aatagcagaa gaccaagaat gggttactct gcttacttat aagcataagc
841 aaaccaaaaa agactataact aagtgtgtt gctttggaaa tatgggtgt tactactttt
901 taaccaaaaa gaaaatgtt accagtccac caagggacgg aggctatccc ttcagtagtg
961 actctggctt gaaaactaac ttgttggaaa aaggcgaacg ccacatgtt agcaactat
1021 atactgatga catcgccca gaaacgggtt agaccacagt aaccacagcg caggaaacta
1081 agcgcggcag aattcaact aagaaggaag tctctattaa gactacactt aaagagctgg
1141 tacataagag agtaacctca ccagaagact ggatgtatgtt gcagccagac agttacattt

1201 aaatgatggc tcaaccaggg ggagaaaaacc tacttaagaa tacgctagag atctgtacgc
 1261 tgactctagc cagaaccaa acagccttg acttgattt agaaaaagct gaaaccagca
 1321 aactaaccaa cttttactg gctgatacaa gaacctgttag aactttgct tttcatggct
 1381 ggaactacat caaagtctgt catgtattt gttgtgtctt gaacagacag ggaggcaaaa
 1441 gaaatactgt tctgtttcat ggaccagcca gtacaggcaa atcaatcatt gcacaggcca
 1501 tagcacaggc agttggtaat gttggttttt ataacgcgc caatgtgaac ttccattta
 1561 atgactgtac caacaagaac ttaatctggg tggagaagaac tggtaacttt ggacagcaag
 1621 taaaccagtt taaagccatt tggctggtc agaccattcg cattgaccaa aaaggaaaag
 1681 gcagcaaca gattgaacca acaccgtga tcattgaccac aaatgaaaac atcacagtgg
 1741 tcaaaatagg gtgtgaagag agaccagaac acactcaacc aatcagagac agaatgttaa
 1801 acattcatct gacacataca ttgcctggc actttggttt gttgataaa aacgaatggc
 1861 ctatgatatg tgcttggcact gtaaagaacg gttaccaatc gaccatggca agttactgt
 1921 ctaaatgggg caaagttcct gattggacag aaaactggc ggagccaaaa gtaacgactg
 1981 aaataaattc ggttagttca accaactcac catctccgaa aagtacgcct ctcagccaga
 2041 actacgcact aactccgtcg gatctcgagg acctggctct ggagccctgg agcacaccaa
 2101 gtactcctgt tgtggcact gtcääacccc cgaacactgg ggaaactgg tcaacagcct
 2161 gtcaagaagc tcaacggagc ccaacttggt ccgagatcga ggaggattt agagcgtgct
 2221 tcagttcggaa acactggaaag agcgactccg aacagctacc aaacttggat taaggtacga
 2281 tggccctcc ggctaaaaga gctaaaagag gtaaggggtt aagggatgg tggtaggtt
 2341 gtggggattt aatatgtgac tacctgtttt acaggcctga aatcacttgg ttctaggtt
 2401 ggtgcctcca ggctacaagt acctgggacc agggAACAGC cttaaccaag gagaaccaac
 2461 caatccatct gacgctgctg ctaaagagca cgacgaggcc tacgaccaat acatcaaatc
 2521 tggaaagaat ccttacccgt acttctctcc tgctgatcaa cgcttcattt accaaaccaa
 2581 agacgctaaa gactggggcg gaaagggttgg tcactacttc tttagaaacca agcgtgctt
 2641 tgcacctaag ctttctactg actctgagcc tgggacttct ggtgtgagca cagctggtaa
 2701 acgtactaaa ccacccgtc acatctttat taaccaagcc agggctaaaa aaaaacgtac
 2761 ttctcttgct ggcgcagcaga ggactcagac aatgagtgtat ggcaccgacc aatctgacag
 2821 cgaaacgct gtccagtcag ctgcttagt tgagcgagca gctgacggc ctggaggc
 2881 tggggccggg ggctctggc ggggtgggt tggctttt actggcagg attataatca
 2941 aacacattat aagtttctag gggatgggtt ggttagagatt actgcttaca gcacacgc
 3001 ggtacacttg aacatgccta aatcagaaaa ctactgtagg gtgcgcgtac acaacacaaa
 3061 tgacacaggt acagcaagtc acatggctat ggacgatgct catgaacaga ttggacacc
 3121 atggagtctg gttgatgcta atgcttggg agttggttt caaccaagtg actggcaga
 3181 catttctaat aatatgattc acatcaattt acattcactt gaccaagaat tggtaatgt
 3241 ggtcatcaaa acagtgtactg aacagaacac aggagctgag gccattaagg tctacaacaa
 3301 tgacccact gctccatga tggctgctct tgattctaac aacatactgc cttacacacc
 3361 agccatagac aatcaagaga cacttggttt ctatccatgg aaaccaacca taccaagtcc
 3421 ttacagatac tatttttagct gtgacagaaaa ctatcagtt acttacaaag acgaaggcagg
 3481 aaccatcaact gacacaatgg gttggccag tggctgaac tcccaatttt ttaccattga
 3541 gaacactcg cgtattaacc tactcagaac tggggatgag tatgctactg gaacttacta
 3601 ctttgacaca gaaccaatca gactaactca cacgtggcaa accaacagac acctgggtca
 3661 gcctccacaa attactgaac taccacgtc tgacactgct aacgctactt taacagctag
 3721 agttacaga tcaggtctga ctcaaattca aggacaaat gatgtgactg aagctactag

3781 ggtcagacct gcacagggtt gattttgtca gcctcatgac aatttgaaa ccagcagagc
3841 ggggccttc aagggtccgg tagtgccagc agacatcaca caaggcctag accatgatgc
3901 caatggtagc ctgagatata cctatgacaa acaacatggt caaagctggg caagtcagaa
3961 caacaaagac aggtacactt gggatgctgt taactatgat tctggcagat ggactaaca
4021 ctgtttattt caatcagtac catttacatc agaaccaa at gctaaccaaa tacttactaa
4081 ccgtgacaac cttagcggta agactgacat acatttacc aacgcattt aca gttatgg
4141 accactaact gc tttccac atcc tgc gcc gatttaccca caaggcaga tttgggacaa
4201 agaacttgat ct tgaacaca agccaagact gcacacacag gtcctttt gtcgtaaaaa
4261 caatgctcca ggtcagcttc tggttaggct agcaccta ac ttgactgacc agt atgatcc
4321 taatagttct aacatctca gaattgtcac ctatggcacc ttcttctgga agggcaact
4381 aactctaaaa gcaaagatga gaccta atgc tacttggAAC ccagtcttcc aaataagtgc
4441 taccaaccaa ggaaccaatg actacatgag cattgaaaga tggttacca ctgctactgg
4501 caacataaca aatgtgcctc tgcttctag acctgttgct agaaacactt actaactaac
4561 tatgctctat gcttcatata tattatatac taactaacca tgttactct
4621 tacattactt catataatata taagactaat aaaaatacaa catagaaata taatattaca
4681 tatagatata aagaatagaa taatatgta cttacttact gtttagaaata atagaacttt
4741 tggataaca agatagttag ttggttatg ttatatagaa tataagaaga tgatgtacaa
4801 agaataaaaag ggtgggaggg tggttgggtg gtactccctt agactgaatg ttagggacca
4861 aaaaaataat aaaattctg aaaacccaaac aaggactact gtcataattca gtgaaccaac
4921 tgaaccatta gtatcaatata tattttaggg tgggggggtg ggagatacat atgttcacta
4981 tggaccaact ggtactgggtt gggtgctctg ctccaaccaa ccagaccggc tctgccggc
5041 tggttgggtt agcgcaacca accagtagtacca gttggtccat agtgaacata tgtatctccc
5101 accccccccac cctaaaaacca tattgataact aatgg

1: NC_001358. Parvovirus H1, co...[gi:9626078]

Links

LOCUS NC_001358 5176 bp ss-DNA linear VRL 20-AUG-2003

DEFINITION Parvovirus H1, complete genome.

ACCESSION NC_001358

VERSION NC_001358.1 GI:9626078

KEYWORDS genome; origin of replication.

SOURCE Parvovirus H1

ORGANISM Parvovirus H1

Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

REFERENCE 1 (bases 1 to 4534)

AUTHORS Rhode,S.L. III and Paradiso,P.R.

TITLE Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybrid-arrested translation

JOURNAL J. Virol. 45 (1), 173-184 (1983)

MEDLINE 83112183

PUBMED 6823009

REFERENCE 2 (bases 4435 to 5176)

AUTHORS Rhode,S.L. III and Klaassen,B.

TITLE DNA sequence of the 5' terminus containing the replication origin of parvovirus replicative form DNA

JOURNAL J. Virol. 41 (3), 990-999 (1982)

MEDLINE 82242308

PUBMED 6284985

COMMENT REVIEWED REFSEQ: This record has been curated by NCBI staff. The reference sequence was derived from X01457.

The viral genome (- strand) is the complementary strand to that shown below (+ strand).

[1] discusses other major open reading frames, but was uncertain as to exact boundaries and/or splicing locations. the non-capsid protein in the features table is speculatively identified as the rf rep gene product: either the postulated site-specific nickase, or the terminal bound protein, or both [1].

ORIGIN

```

1 cattttaga actgaccaac catgttcacg caagtgcacgt gatgacgcgc gctgcgcgcg
61 ctgccttcgg cagtcacacg tcacttagcgt ttcacatgggt tggcagttc taaaaatgat
121 aagcggttca gagagttga aaccaaggcg ggaaacggaa gtgggcgtgg ctaactgtat
181 ataagcagtc actctggtcg gttactcaact ctgcattcat ttctgagttt gtgagacaca
241 ggagcgagac taaccaacta accatggctg gaaacgctta ctccgatgag gttttgggag
301 taacaaactg gctgaaggac aaaagttagcc aggaggtttt ctcattttgtt tttaaaaatg

```

361 aaaacgtcca actaaatgga aaggacatcg gttggaatag ttacagaaaag gagctacaag
421 atgacgagct gaagtctcta caacgagggg cggagaccac ttgggaccaa agcaggaca
481 tggaaatggga gagcgcagtg gatgacatga ccaaaaagca agtattttt tttgattctt
541 tggtaagaa gtgttgttt gaagtgccta gcacaaagaa catagtcct agtaatgtta
601 ctgggtcgt gcagcatgaa tggggaaagg acccaggctg gcactgtcat gtgctgattg
661 gaggcaagga cttagtcaa cctcaaggaa aatggtgag aaggcagcta aatgtgtact
721 ggagtagatg gttgtgact gcctgtaatg tcactaactac accagctgaa agaattaaac
781 tgagagaat agcagaggac agtgaatggg tcacttgct tacctataag cataagcaca
841 ccaagaagga ctataccaag tgtgttctt ttggaaacat gattgttat tacttttaa
901 gaaaaaaagaa aatatgtacc agtccaccaa gggacggagg ctatttctt agcagtgact
961 ctggctggaa aactaacttt ttgaaagagg gcgagcgcctt tctagtgagc aaactgtata
1021 ctgatgagat gaaaccagaa acggcgtgaga ccacagtgc cactgcacag gaagctaagc
1081 gccccagaat tcaaactaga gaggaggtct cgattaaaac cacactcaaa gagttggcac
1141 ataaaagagt aacccacca gaagactgga tcatgtatgc gcccacatg tacattgaaa
1201 tcatggctca accaggtgga gaaaacttgc taaaaatac actagagatc tgcactgt
1261 ctctagcaag aaccaaaaca gccttgact tgattctgga aaaagctgaa accagcaaac
1321 tagccaaactt ttccatggct agcaccagaa cctgtagaat ctggctgag catggctgga
1381 actatattaa agtctgcctt ggcattgtt gtgtgtgaa tagacaagga ggcaaaagga
1441 acactgtgct ctgcacgca ccagccagca caggccaaatc tattattgca caagccatag
1501 cacaaggat tggtaatgtt ggttgtaca atgctgcca tggtaacttt ccatttaatg
1561 actgtaccaa caaaaacttg atttgggtgg aagaagctgg taactttggc cagcaagtaa
1621 accaattcaa agctattgt tctggccaaa ccatacgc tgcataaaaaa ggaaaaggca
1681 gcaaaacatg tgaaccaaca ccagtttata tgaccaccaa cgagaacatt accgtggta
1741 gaataggctg tgaggaaaga ccagaacaca ctcaccaat cagagacaga atgctcaaca
1801 ttccacctgac acgtacacta cctggtgact ttggtttgtt ggataagcac gaatggcctc
1861 tgcattgtgc ttgggttgtt aagaatgtt accaatctac catggctgt tactgtgcta
1921 aatggggcaa agttcctgat tggtcagagg actggggcggg gcccggacta gacactccta
1981 taaattcgct aggtcaatg cgctcaccat ctgtactcc gagaagtcg cctctcagcc
2041 aaaactacgc tcctactcca ctgcacatgg accttgcggg cctagctcta gggccctgga
2101 gcacaccaaa tactcctgtt gccccactg cggcaagccaa acacactggg gaggctgggt
2161 ccacagcctg ccaagggtgct caacggagcc caacctggc cgagatcgag gggatttga
2221 gagcttgctt cagtcaagaa cagttggaga ggcacatcaa cgaggagctg accttggact
2281 aaggtacaat ggcacactca gctaaaagag ctaaaagagg taagggctt agggatgggt
2341 ggttgtgggg gtactaatgt atgactaccc ttgttacagg cctgaaatca ttgggtctt
2401 ggttgtgggc tcctctggctt caagtacactg ggaccaggaa acggccctt gcaaggagaa
2461 ccaaccaacc ttctgcacgc cgctgcaaaa gaacacgacg aacccatcg ccaatacatc
2521 aaatctggaa aaaatccctt cctgtacttc tctcctgctt atcaacgctt cattgaccaa
2581 accaaagacg ccaaggactg gggccggcaag gttggactt acttttttag aaccaagcga
2641 gctttgcac ctaagcttc tactgactt gaaacctggca cttctgggtt gagcagaccc
2701 ggttaacgaa ctaaaccacc tgctcacatt ttgttaatc aagccagagc taaaaaaaaa
2761 cggcttc tgcgtgcaca gcagaggact ctgacaatga gtgtggcac cggaaacaaac
2821 caaccagaca ctgaaatcgc taatgtt gttggactt cagctgacgg aggtggaaagc
2881 tctgggggtt gggctctgg cgggggtggg attgggtttt ctactggac ttatgtataat

2941 caaacgactt ataagtttt gggagatgga tggtagaaaa taactgcaca tgcttcaga
3001 ctttgcact tggaatgcc tccttcagaa aactactgcc gcgtcaccgt tcacaataat
3061 caaaacaacag gacacggAAC taaggtaaAG gggAAACATGG CCTATGACAC acatcaacaa
3121 attggcacAC catggagCTT ggttagatGCT aatgttggg gagtttgggtt ccaaccaAGT
3181 gactggcagt tcattcaaaa cagcatggaa tcgctgaatC ttgactcatt gagccaAGAA
3241 ctatTTAATG tagtagtcaa aacagtcaCT gaacaacaAG gagctggccA agatGCCATT
3301 aaagtctata ataatgactt gacggcCTGT atgatggTT ctctggatAG taacaacata
3361 ctgccttaca cacctgcAGC tcaaACATCA gaaACACTTG gtttctACCC atggAAACCA
3421 accgcaccAG ctccttacAG atactacttT ttcatGCCTA gacaACTCAG tgtaacCTCT
3481 agcaactCTG ctGAAGGAAAC tcaaATCACA gacACCATTG gagAGCCACA ggcactAAAC
3541 tctcaatttt ttactattGA gaacACCCTG CCTATTACTC tcctgcgcAC aggtgatgag
3601 ttTACAACTG gcacccTACAT CTTAACACT gacCCACTA aactactca cacatggCAA
3661 accaACAGAC acTTGGCATG CCTCCAAGGA ataactgacc taccaACATC agatacAGCA
3721 acagcatCAC taactgCAA TGGAGACAGA ttggatCAA cacAAACACA gaatgtGAAC
3781 tatgtcacAG aggcttgcG caccAGGCT gctcAGATTG gcttcatGCA acCTcatGAC
3841 aactttGAAG caaacAGAGG tggccATT aaggTTCCAG tggTACCGCT agacataACA
3901 gctggcgagg accatgatGC aaacGGAGCC atacgattTA actatggCAA acaACATGGC
3961 gaagattggg ccaaacaAGG agcAGCacCA gaaAGGTACA catgggatGC aattgatAGT
4021 gcagctggGA gggacacAGC tagatGTT gtacAAAGTG caccaATATC tattccACCA
4081 aaccaAAACC agatCTTGCA gcgagaAGAC gccatAGCTG gcagaACTAA catgcATTAT
4141 actaatgttt ttaacagCTA tggTCCACTT agtgcattTC ctcatCCAGA tcccatttat
4201 ccaaATGGAC aaatttggGA caaAGAAATTG gacCTGGAAc acaaACCTAG actacACGTA
4261 actgcaccat ttgttGtaa aaacaACCCa ccaggTcaac tatttgtca ctggggcCT
4321 aactCTGACTG accaATTGA cccaaACAGC acaACTGTTT ctCGCATTGT tacatATAGC
4381 acttttact ggaagggtat ttGAAATTc aaAGCCAAAC taAGACCAAA tctgacCTGG
4441 aactCTGTat accaAGCAAC cacAGACTCT gttGCAATTt ctTACATGAA tGTTAAGAAA
4501 tggctcccat ctgcaactgg caacatgcAC tctgatCCAT tgatttGTag acctgtGCC
4561 cacatgacat actAACCAAC caactatGTT tctctgttG cttcacataa tacttaaACT
4621 aactagacta caacataAAA atatacactt aataatAGAT tattaaaat aacataataat
4681 ggtaggTTA ctgtaaaaaaa taatAGAACT ttGGAATAA atatAGTTAG ttggTTAATG
4741 ttagatGAA tataAAAGA ttttgatTT taaaataAA atatGTTAGT ggttaatGTT
4801 agatAGAA taaaAGATT ttgtatttgg gaaataAAAAA gggtggTTgg ttggTTggTT
4861 ggtactccCT tagactGAAT gttAGGGACC AAAAATAA taaaATAATT AAAATGAACA
4921 aggactactG tctattcAGT tgaccaACTG AACCTATAGT ATCACTATGT ttttagggTG
4981 ggggggtggg agatacatac gttcgctatG gaccaAGTGG taccGGTGG ttGCTAAGCT
5041 cgaacaAGAC ggctaAGCCG gtccGGTTGG ttGAGCGCAA ccaACCGGTa ccacttggTC
5101 catagcgaAC gtatgtatCT cccACCCCCC caccCTAAAC acatAGTGT acTATAGTT
5161 cagtggTCA actgaa